U.S. Environmental Protection Agency
Information Resources Management (IRM) Strategic Plan

The collection, use, and dissemination of information of known and appropriate quality are integral to ensuring that the U.S. Environmental Protection Agency (EPA) achieves its mission. Information about human health and the environment -- environmental characteristics; physical, chemical, and biological processes; and chemical and other pollutants -- underlies all environmental management and health protection decisions. The availability of, and access to, information and the analytical tools to understand it are essential for assessing environmental and human health risks, designing appropriate and cost-effective policies and response strategies, and measuring environmental improvements. For these reasons the Agency is focusing important attention on Information Resources Management (IRM). This IRM Strategic Plan outlines the agency's vision for ensuring the appropriate management of information to ensure the greatest utilization by the Agency for optimum outcomes while at the same time providing environmental information and developing important information tools to better enable the public to actively participate in environmental decisions in their communities and on a national scale.

This document represents an ongoing commitment by EPA to focus on and improve EPA's overall information management and information technology (IM/IT) programs in order to achieve the Agency's mission to protect human health and the environment. In doing so, this document is intended to communicate two main concepts. First this plan will become a home for describing how EPA is meeting relevant information dissemination requirements, the results of our program review process and our plans to reduce performance gaps we may identify. And second, this plan will also describe the Agency's information models in place and how those models relate to the overarching Data Reference Model (DRM).

Organizing and Categorizing Public Access Information

An essential component of the Agency's information management strategy is its web site, www.epa.gov. It is the Agency's primary public interface. Its pages are a fundamental part of every Agency program; taken together, those pages are the foundation of the Agency's environmental outreach program. The site conveys and amplifies the Agency's mission, goals and work. As such, priority is given to making environmental information publicly available, accessible, and searchable.

The U.S. Office of Management and Budget (OMB) has published guidance calling for federal agencies to "Organize and categorize information intended for public access and ensure it is searchable across agencies." EPA is doing many things to accomplish these goals, and the Agency has projects underway that will improve its ability to meet these requirements. Specifically, EPA is currently:

- Using a new search engine on its Web site. It can deliver results as XML, so EPA is poised to respond to "federated" searches when requested. It is metadata-

enhanced and has a rich feature set for advanced searches and for limiting searches to sub-areas of EPA's Web site.

- Using a metadata catalog to enhance search results and to provide a key browse capability (at http://www.epa.gov/epahome/topics.html). The metadata catalog is based on a controlled vocabulary of topics, terms, and keywords assigned to the most important Web pages.
- Restructuring the Web contents to use a new template and Cascading Style Sheets for better control over look-and-feel and other enhancements. The new template includes ICGI-identified metadata elements in each page header. The new template is currently being implemented with a projected completion date of August 2007.
- Providing a search engine on the Agency's information resource inventory to discover those systems that contain publicly accessible data: www.epa.gov/read

Additional projects which we intend to complete in the next year include:

- Deploying a new multi-faceted taxonomy to replace the current topics-only controlled vocabulary. This will further enhance EPA's ability to drive visitors to the specific information they are seeking. (December 2006)
- Enhancing the search engine to include additional browse capabilities through topical folders.
- Completing a Web Content Management Pilot whose success will provide a greater ability to acquire and use metadata; to deploy the new information architecture; to cross-link information on the site; to render it in a variety of formats (e.g., XML, mobile-access, screen-reader access-specific); and to reduce redundancy and improve consistency across the site. (March 2007)
- Improving Web site traffic analysis using performance metrics and web analytical tools. (estimated March 2007)

EPA's Use of the Federal Enterprise Architecture Data Reference Model (DRM)

The EPA Enterprise Architecture (EA) program has developed a new EA metamodel to describe the total set of objects, properties, and relationships available for use in modeling and defining architectures for the Agency. This metamodel has been implemented in Metis, the Agency EA standard tool, as Architecture Repository and Tool (ART) 4.0.

When the metamodel was developed, the EA program also recognized that representing interfaces and data exchanges with other federal agencies, state governments, industry, academia and tribes would be essential to characterize a variety of interoperability and data sharing issues. For those purposes, a number of Partner Architecture elements were developed to provide architects the ability to model and represent touch-points that EPA has with these outside organizations and the information being exchanged.

Focusing on the data domain of the EPA metamodel, the abstract model of the FEA DRM Draft version 2.0, November 17, 2005 was the blueprint for identifying the metamodel elements and relationships that now drive development and representation of the data layer of the EPA Enterprise Architecture, as well as the data layers of the constituent EPA segment architectures. The three FEA DRM standardization areas, *Data Sharing*, *Data Description* and *Data Context,* are represented directly and explicitly in the EPA EA metamodel.

Registry of EPA Applications and Databases

The EPA's Registry of Applications and Databases (READ) is the authoritative registry for uniquely identifying EPA's information resources (e.g., data systems and models). A comprehensive inventory is necessary for the Agency to manage effectively its diverse holdings. The OMB has directed agencies to establish system inventories to improve information management.

For each of the approximately 1500 information resource records in READ, there is basic information, including the title, acronym, description, the statute supported by the system, how it fits into the Agency's strategic plan, and whether the data is accessible to the public. READ also includes fields for the system platform, which geospatial tools are used, the lifecycle phase, and identifies other information resources for which there is an interface or other relationship. The public version of READ is available at: http://www.epa.gov/read.

Upcoming enhancements planned for October 2006 will make READ easier to use and integrates it with the other information management tools on capital planning and investment, security, and privacy. READ also will be used as the tool for collecting enterprise architecture information, which then will be exported on a nightly basis to the Architecture Repository Tool (ART).